Discrete Mathematics 82 (1990) 1–6
North-Holland

INVITED PAPER

GENERALIZED DYCK PATHS*

Jacques LABELLE
Université du Québec à Montréal, Département de Mathématique et l'Informatique, Case postale 8888, Succ. "A", Montréal, Qué., H3C 3P8, Canada

Yeong-Nan YEH
Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan, R.O.C.

Communicated by G.-C. Rota

Received 11 March 1988

It is well known (see [3, 6, 9, 10, 11]) that Dyck paths are in bijection with "Dyck words", "ballot sequences", "well formed sequences of parentheses", "2-lines standard-tableaux", "binary trees", "ordered trees"; all these are counted by Catalan numbers. In the present text, we replace the north-east steps of a Dyck path by steps from an arbitrary finite multi-set $\mathcal{S}$ of vectors with integral coordinates in the plane. In order to study these generalized Dyck paths, called A-paths, we have to introduce many closely related families of paths. The corresponding (multi-variable) generating functions satisfy an intricate system of algebraic equations which leads to a polynomial equation satisfied by $\boldsymbol{A}$. For example when $\mathcal{S}$ is $\{(1,1)\}$ (respectively $\{(1,2),(2,1)\}$; $\{(1,3),(3,1)\}$) this polynomial equation is of degree 2 (resp. 4, 8). More generally when $\mathcal{S}=\{\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_m\}$ where $\boldsymbol{u}_j=(r_j,j)$ then $\boldsymbol{A}=\boldsymbol{A}(u_1,u_2,\ldots,u_m)$ satisfies a polynomial equation of degree 2^m with coefficients in $\mathbb{Z}[u_1,u_2,\ldots,u_m]$.

0. Introduction

Let $\mathcal{S}$ be a finite multi-set (i.e. a set with repetitions) of vectors (or steps) in $\mathbb{N}\times\mathbb{N}$ where $\mathbb{N}=\{1,2,\ldots,\}$. For $\boldsymbol{u}=(r,s)\in\mathcal{S}$, we write $\boldsymbol{u}^*=(r,-s)$, $r=\pi_1(\boldsymbol{u})$ and $s=\pi_2(\boldsymbol{u})$. Repetitions of the same vector in $\mathcal{S}$ may be thought as steps of different colors.

Definition 1. An $\mathcal{S}$-Dyck path (for short we call these A-paths) is a path in $\mathbb{Z}\times\mathbb{Z}$ which:
(a) is made only of steps in $\mathcal{S}+\mathcal{S}^*$
(b) starts at $(0,0)$ and ends on the x-axis
(c) never goes strictly below the x-axis.
If it is made of l steps and ends at $(n,0)$, we say that it is of length l and size n.

Definition 2. For $k\geqslant 0$, an $A^{(k)}$-path is an A-path which touches the x-axis, not counting at $(0,0)$, exactly k times.

* Partially supported by grants NSERC OGP0037460 (Canada) and FCAR EQ 1608 (Québec).

Definition 3. For $i>0$, a B_i-path is a path in $\mathbb{Z}\times\mathbb{Z}$ which satisfies (a), (b) and (c) of Definition 1 except that it ends at (n, i) for some n. If it touches the x-axis, not counting at $(0,0)$, exactly k times, we call it a $B_i^{(k)}$-path. We set $B_0=A$.

Definition 4. For $i>0$, a C_i-path is a path in $\mathbb{Z}\times\mathbb{Z}$ which satisfies (a), (b) and (c) of Definition 1 except it needs only stay above the line $u=-i$. If it touches this line exactly k times, we call it a $C_i^{(k)}$-paths. We set $C_0=A$.

Let $\mathcal{S}=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ and $u_1, u_2, \ldots, u_m$ be corresponding formal variables. Consider the m-variables generating function

$$\boldsymbol{A}=\boldsymbol{A}(u_1, u_2, \ldots, u_m)=\sum_{i_j\geqslant 0} a_{i_1 i_2\cdots i_m} u_1^{i_1} u_2^{i_2}\cdots u_m^{i_m}$$

where $a_{i_1 i_2\cdots i_m}$ is the number of A-paths in which $\forall 1\leqslant j\leqslant m$, $\boldsymbol{u}_j$ and $\boldsymbol{u}_j^*$ appear a total of i_j times. Similarly we define the generating functions $\boldsymbol{A}^{(k)}$, $\boldsymbol{B}_i$, $\boldsymbol{B}_i^{(k)}$, $\boldsymbol{C}_i$ and $\boldsymbol{C}_i^{(k)}$.

Lemma 1. *If we set $u_j=t$ (respectively $u_j=t^{r_j}$; $u_j=xy^{r_j}$ where $r_j=\pi_1(\boldsymbol{u}_j)$) in $\boldsymbol{A}(u_1, u_2, \ldots, u_m)$, we obtain the generating function for A-paths with respect to length (respectively size; length in x and size in y) and similarly for the other generating series.*

Main Theorem. *Let $\mathcal{S}=\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m\}$ where $\boldsymbol{u}_j=(r_j, j)$, $1\leqslant j\leqslant m$, $r_j>0$. The generating functions $\boldsymbol{A}$, $\boldsymbol{A}^{(1)}$, and for $1\leqslant i<m$, $\boldsymbol{B}_i$, $\boldsymbol{B}_i^{(0)}$, $\boldsymbol{C}_i$, $\boldsymbol{C}_i^{(0)}$ satisfy the following system of $4m-2$ algebraic equations (recall that $\boldsymbol{B}_0=\boldsymbol{C}_0=\boldsymbol{A}$):*

(1) $\boldsymbol{A}\cdot\boldsymbol{A}^{(1)}+1=\boldsymbol{A}$

(2) $\boldsymbol{A}^{(1)}=\sum_{j=1}^{m}\boldsymbol{C}_{j-1}u_j^2+2\sum_{1\leqslant i<j\leqslant m}\left(\sum_{k=1}^{i}\boldsymbol{B}_{j-k}^{(0)}\cdot\boldsymbol{B}_{i-k}\right)u_i u_j$

(3) $\boldsymbol{B}_i=\boldsymbol{A}\cdot\boldsymbol{B}_i^{(0)}$ for $1\leqslant i<m$

(4) $\boldsymbol{C}_i=\boldsymbol{C}_{i-1}+\boldsymbol{B}_i^{(0)}\cdot\boldsymbol{B}_i$ for $1\leqslant i<m$

(5) $\boldsymbol{C}_i^{(0)}=\boldsymbol{C}_{i-1}$ for $1\leqslant i<m$

(6) $\boldsymbol{B}_i^{(0)}=\sum_{j=1}^{m}\left(\sum_{k=1}^{\min(i,j)}\boldsymbol{B}_{j-k}^{(0)}\cdot\boldsymbol{B}_{i-k}\right)u_j$ for $1\leqslant i<m$.

Proof. (1) $\boldsymbol{A}=\sum_{k\geqslant 0}\boldsymbol{A}^{(k)}=\sum_{k\geqslant 0}(\boldsymbol{A}^{(1)})^k=(1-\boldsymbol{A}^{(1)})^{-1}$.

(2) In an arbitrary $A^{(1)}$-path, erase the first and last steps. If these steps are $\boldsymbol{u}_i$ and $\boldsymbol{u}_i^*$ one gets a C_{i-1}-path; if they are $\boldsymbol{u}_i$ and $\boldsymbol{u}_j^*$ (with $1\leqslant i<j<m$) one gets, after translation, a path from $(0,0)$ to $(n-r_i-r_j, j-i)$ which always stays above the line $y=-i+1$. These paths have generating function

$$\sum_{k=1}^{i}\boldsymbol{B}_{j-k}^{(0)}\cdot\boldsymbol{B}_{i-k}$$

where $y=k-i$ is the lowest horizontal line which the path touches.

(3) easy;

(4) easy;

(5) trivial; in fact we may very well eliminate the $C_i^{(0)}$'s and equation (5) from the system.

(6) Removing the first step (say u_j) of a $B_i^{(0)}$-path gives a path from (r_j, j) to (n, i) for some n which always stays above the line $y = 1$. After translation this corresponds to a path from $(0, 0)$ to $(n - r_j, i - j)$ which stays above $y = 1 - j$. These paths have generating function $\sum_{k=1}^{\min(i,j)} B_{j-k}^{(0)} B_{i-k}$ as was proved in (2). □

Theorem 1. *The generating functions A, B_i, $1 \leqslant i < m$, satisfy the following system of m equations (where $A = B_0$ and $B_{-n} = 0$ for $n > 0$):*

$$(*) \qquad A = 1 + \sum_{1 \leqslant i,j \leqslant m} \Big(\sum_{k \geqslant 1} B_{i-k} B_{j-k} \Big) u_i u_j$$

$$(**) \qquad B_i = \sum_{j=1}^{m} \Big(\sum_{k \geqslant 1} B_{i-k} B_{j-k} \Big) u_j \quad \text{for } 1 \leqslant i < m.$$

Proof. To obtain $(*)$; multiply (2) by A, substitute in (1), use (3) and (4′) $AC_i = AC_{i-1} + B_i^2$. To obtain $(**)$; multiply (6) by A. □

Theorem 2. *Let $\mathcal{S}$ be arbitrary finite multi-set and $m = \max_{u \in \mathcal{S}} \pi_2(u)$. If in the algebraic system of Theorem 1, we replace $u_j (1 \leqslant j \leqslant m)$ by $\sum_{\pi_2(u)=j} u$ (note that this sum may very well be zero), we get an algebraic system of m equations satisfied by the $|\mathcal{S}|$-variables generating functions A, B_i, $1 \leqslant i < m$.*

Remark 1. We could also add "horizontal steps" in our generalized Dyck paths (these could be called generalized Motzkin paths): let $u_0 = (r_0, 0)$, $r_0 > 0$, and u_0 be the corresponding variable. The system of equations satisfied by the $(m+1)$-variables generating functions A, $A^{(1)}$, B_i, $B_i^{(0)}$, C_i and $C_i^{(0)}$ $(1 \leqslant i < m)$ is the same except the extra term u_0 is added to the right hand side of (2). For example, taking $\mathcal{S} = \{(1, 0), (1, 1)\}$ leads to Motzkin paths (with bicolored horizontal steps) whose generating function (with respect to length = size) $M(t)$ satisfies:

$$t^2 M + (2t - 1)M + 1 = 0 \quad \text{and } M = \frac{1}{2t^2}\big(1 - 2t - \sqrt{1 - 4t}\big).$$

Recall that there is a nice bijection [11] between (classical) Dyck paths of length $2n + 2$ and bicolored Motzkin paths of length n (both of these are counted by the $(n + 1)$th-Catalan number): after erasing first and last steps of the Dyck path replace successively the "pics", "creux", "doubles montées" and "doubles descentes" by respectively "horizontal steps of the first color", "of the second color", "montées" and "descentes" to get the corresponding bicolored Motzkin path.

1. $|\mathscr{S}| = 1$ (classical Dyck paths)

When $\mathscr{S} = \{\boldsymbol{u}\}$, $\boldsymbol{u} = (r, 1)$, the system in the main theorem becomes (with $u_1 = u$):

(1) $\boldsymbol{A} = \boldsymbol{A} \cdot \boldsymbol{A}^{(1)} + 1$; (2) $\boldsymbol{A}^{(1)} = u^2 \boldsymbol{A}$

Lemma 2. *We have* $u^2\boldsymbol{A}^2 - \boldsymbol{A} + 1 = 0$ *and* $\boldsymbol{A} = (1/2u^2)(1 - \sqrt{1 - 4u^2})$. *Moreover we have* $a_{2m} = (1/m + 1)\binom{2m}{m}$ *(the* m*th-Catalan number) and* $a_{2m+1} = 0$.

Remark 2. In general for $\boldsymbol{u} = (r, s)$, $s > 0$; $\boldsymbol{A}$ is the same (just shrink everything by a factor $1/s$ in the y direction). The case $\boldsymbol{u} = (r, 0)$ is different: $\boldsymbol{A} = (1 - 2u)^{-1}$; $a_n = 2^n$.

2. $|\mathscr{S}| = 2$

When $\mathscr{S} = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, $\boldsymbol{u}_1 = (r_1, 1)$, $\boldsymbol{u}_2 = (r_2, 2)$, we get from Theorem 1 (with $u_1 = u$, $u_2 = v$) the following 4th-degree polynomial equation satisfied by $\boldsymbol{A}$:

Proposition 1. *The generating function* $\boldsymbol{A} = \boldsymbol{A}(u, v)$ *satisfies the polynomial equation*:

$$v^4\boldsymbol{A}^4 - v^2(2v + 1)\boldsymbol{A}^3 + (u^2 + 2v^2 + 2v)\boldsymbol{A}^2 - (2v + 1)\boldsymbol{A} + \mathbf{1} = \mathbf{0}.$$

Using MAPLE one then obtains the explicit expression:

$$\boldsymbol{A}(u, v) = \frac{1}{4u^2}\{1 + 2u + \Delta - \sqrt{2}\sqrt{1 - 4u^2 - 2v^2 + (2u + 1)\Delta}\},$$

where

$$\Delta = \sqrt{4u^2 - 4u + 1 - 4v^2}.$$

Setting $u = t^2$ and $v = t$ (resp. $u = v = t$) the polynomial equation and explicite expression for the generating series of "Dyck paths of chess-knight moves with respect to size (resp. length)", which was found in [8], is obtained.

3. $|\mathscr{S}| = 3$

When $\mathscr{S} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, $\boldsymbol{u}_i = (r_i, i)$, $1 \leqslant i \leqslant 3$, we get from solving the system of Theorem 1 (with $u_1 = u$, $u_2 = v$, $u_3 = w$):

Proposition 2. *The generating function* $\mathbf{A} = \mathbf{A}(u, v, w)$ *satisfies the polynomial equation*:

$$\begin{aligned} & w^8\mathbf{A}^8 - w^6(2v+1)\mathbf{A}^7 + w^4(u^2+2v^2+2v-2uw+w^2)\mathbf{A}^6 \\ &\quad - w^2(2v^3+v^2-4uvw-2uw+2vw^2+w^2)\mathbf{A}^5 \\ &\quad + (v^4+2u^2w^2-4uv^2w-4uw^3+4v^2w^2+w^2)\mathbf{A}^4 \\ &\quad - (2v^3+v^2-4uvw-2uw+2vw^2+w^2)\mathbf{A}^3 \\ &\quad + (u^2+2v^2+2v-2uw+w^2)\mathbf{A}^2 - (2v+1)\mathbf{A} + \mathbf{1} = \mathbf{0} \end{aligned}$$

Remark 3. This can also be written:

$$\begin{aligned} & (v^4+2u^2w^2-4uv^2w-4uw^3+4v^2w^2+w^2)\mathbf{A}^4 \\ &\quad - (1+w^2\mathbf{A}^2)(2v^3+v^2-4uvw-2uw+2vw^2+w^2)\mathbf{A}^3 \\ & (1+w^4\mathbf{A}^4)(u^2+2v^2+2v-2uw+w^2)\mathbf{A}^2 - (1+w^6\mathbf{A}^6)(2v+1)\mathbf{A} \\ &\quad + (1+w^8\mathbf{A}^8) = \mathbf{0}. \end{aligned}$$

Of course, setting $w = 0$ gives back Proposition 1; setting $u = v = 0$ gives a polynomial which is divisible by $w^2\mathbf{A}^2 - \mathbf{A} + 1$ (see Remark 2 and Lemma 2).

Remark 4. In general, when $\mathscr{S} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ where $\mathbf{u}_j = (r_j, j)$, $\mathbf{A}(u_1, u_2, \ldots, u_m)$ seems to satisfy a polynomial equation of degree 2^m with leading coefficient $u_m^{2^m}$.

Corollary 1. *Setting* $u = t^3$, $v = 0$ *and* $w = t$, *in Proposition* 2, *we find that* $\mathbf{A}(t)$, *the generating series for Dyck paths of* "(1, 3)-*knight moves*" *with respect to size, satisfies*:

$$\begin{aligned} & t^8\mathbf{A}^8 - t^6\mathbf{A}^7 + (t^{10}-2t^8+t^6)\mathbf{A}^6 + (2t^6-t^4)\mathbf{A}^5 \\ &\quad + (2t^8-4t^6+t^2)\mathbf{A}^4 + (2t^4-t^2)\mathbf{A}^3 + (t^6-2t^4+t^2)\mathbf{A}^2 - \mathbf{A} + \mathbf{1} = \mathbf{0}. \end{aligned}$$

Corollary 2. *Setting* $u = 0$, $v = t^3$ *and* $w = t^2$, *in Proposition* 2, *we find that* $\mathbf{A}(t)$, *the generating series for Dyck paths of* "(2, 3)-*knight*" *moves with respect to size, satisfies*:

$$\begin{aligned} & t^{16}\mathbf{A}^8 - (2t^{15}+t^{12})\mathbf{A}^7 + (2t^{14}+t^{12}+2t^{11})\mathbf{A}^6 - (2t^{13}+2t^{11}+t^{10}+t^8)\mathbf{A}^5 \\ &\quad + (t^{12}+4t^{10}+t^4)\mathbf{A}^4 - (2t^9+2t^7+t^6+t^4)\mathbf{A}^3 \\ &\quad + (2t^6+t^4+2t^3)\mathbf{A}^2 - (2t^3+1)\mathbf{A} + \mathbf{1} = \mathbf{0} \end{aligned}$$

Corollary 3. *Setting* $u = v = w = t$, *in Proposition* 2, *we find that* $\mathbf{A}(t)$, *the generating series with respect to length for Dyck paths with* $\mathscr{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$, *satisfies*:

$$\begin{aligned} & t^8\mathbf{A}^8 - t^6(2t+1)\mathbf{A}^7 + 2t^5(t+1)\mathbf{A}^6 - t^2(t^2-1)\mathbf{A}^4 \\ &\quad + 2t(t+1)\mathbf{A}^2 - (2t+1)\mathbf{A} + \mathbf{1} = \mathbf{0}. \end{aligned}$$

Remark 5. We easily get a recurrence formula for the coefficients of the series $A(u_1, u_2, \ldots, u_m)$ from the polynomial equation it satisfies. For example, Lemma 2 gives $a_0 = 1$, $a_{n+1} = \sum_{i+j=n} a_i a_j$.

Conclusion. We may think of all these A-paths as words in the letters x, y, $\bar{y}$ (by replacing $\boldsymbol{u} = (r, s)$ (resp. $\boldsymbol{u}^* = (r, -s)$) by the word $x^r y^s$ (resp. $x^r \bar{y}^s$). The corresponding languages are algebraic but the degree of the polynomial equation they satisfy becomes rapidly very high. A closely related subject which is certainly worth studying is "Brownian motions in the plane" formed with steps (r, s), $(r, -s)$, $(-r, s)$, $(-r, -s)$ where $\boldsymbol{u} = (r, s)$ is taken from a given finite multi-set of vectors.

Acknowledgements

The authors wish to thank Germain Kreweras, for the very inspiring talk [6] and encouragements, François Bergeron for some good suggestions and Ivan Constantineau for doing the computer work.

References

[1] C. Berge, Graphes et hypergraphes (Dunod, 1973).
[2] L. Comtet, Analyse Combinatoire (Presses Universitaires de France, SUP, 1970).
[3] I.P. Goulden and D.M. Jackson, Combinatorial Enumeration (John Wiley and Sons, 1983).
[4] A. Kotzig and J. Labelle, Sur les graphes de mouvements de cavaliers généralisés, Centre de Recherche en Mathématiques, Univ. de Montréal, Rapport no. 1460.
[5] M. Kraitchik, Recreationnal Mathematics (Dover, 1942).
[6] G. Kreweras, Combinatoire énumérative des ponts de Catalan (ou mots de Dyck), Talk at UQAM (April 16th 1987).
[7] J. Labelle, Théorie des Graphes (Modulo Editeur, 1981).
[8] J. Labelle and Y.N. Yeh, Dyck paths of knight moves, Discrete Appl. Math. 24 (1989) 213–221.
[9] R. Stanley, Enumerative Combinatorics; Vol. 1 (Addison-Wesley, 1986).
[10] D. Stanton and D. White, Constructive Combinatorics (U.T.M., Springer Verlag, 1986).
[11] G.X. Viennot, Théorie combinatoire des nombres d'Euler et de Genocchi, Séminaire de théorie des nombres, Publications Univ. Bordeaux I, année 1981/82.